JITTA

JOURNAL OF INFORMATION TECHNOLOGY THEORY AND APPLICATION

A Publication of the Association for Information Systems

JOURNAL OF INFORMATION TECHNOLOGY THEORY AND APPLICATION

A Content Validity Study for a Knowledge Management Systems Success Model in Healthcare

Nor'ashikin Ali
Department of Information Systems
Universiti Tenaga Nasional, Kajang, Malaysia.
shikin@uniten.edu.my

Alexei Tretiakov
School of Management
Massey University, Palmerston North, New Zealand
alexei.tretiakov@gmail.com

Dick Whiddett
School of Management
Massey University, Palmerston North, New Zealand
R.J.Whiddett@massey.ac.nz

Abstract:

Although research in knowledge management systems (KMS) has been growing, its instrument development has not been given sufficient attention. In particular, the issue of content validity has rarely been addressed in a systematic manner. Formally demonstrating content validity is an important step in the instrument-development process, and relies on a panel of experts' judgment to confirm that the items generated are congruent with the constructs. This papers reports on a content validity study for a KMS success model in healthcare. The study procedures involved selecting experts and collecting and analyzing their evaluations. We formed the panel of experts by contacting the participants of a conference on health informatics and the authors of papers related to knowledge management published in major journals. We used Lawshe's (1975) technique, which uses the computation of content validity ratio (CVR) to screen questionnaire items. This study will help practitioners and researchers in KMS to create valid and reliable measure.

Keywords: Content Validity, KMS, CVR, Lawshe's Technique, Healthcare.

Volume 15, Issue 2, pp. 21-36, June 2014

Marcus Rothenberger acted as the Senior Editor for this paper.

A Content Validity Study for a Knowledge Management Systems Success Model in Healthcare

INTRODUCTION

As more organizations attempt to implement knowledge management systems (KMS), researchers and practitioners are increasingly interested in the factors that influence their successful adoption. Several conceptual models for KMS success that attempt to identify the factors that influence an implementation's success have been developed (e.g., Halawi, McCarthy, & Aronson, 2007; Jennex & Olfman, 2005; Kulkarni, Ravindran, & Freeze, 2007; Wu & Wang, 2006). Most prior research of KMS success has been conducted in business organizations; however, because the healthcare industry is highly knowledge intensive, there is an increasing interest in using KMS to help healthcare practitioners access up-to-date medical knowledge and share their knowledge with each other (Winkelman & Choo, 2003). Healthcare organizations differ from organizations in other industries for reasons such as the need to keep up with constantly evolving medical knowledge while maintaining high ethical standards. Therefore, study factors contributing to KMS success—specifically in the healthcare context—is important. Although prior studies influenced our research (e.g., we often adapted existing research instruments for use in the questionnaire), we revalidated the instruments to ensure that they were still applicable in the healthcare context.

In this paper, we discuss the various approaches to determining the content validity of survey instruments and describe the process we followed in our research. We then report on the findings for the validity of a range of instruments commonly used in research; in this way, we place the previous researchers' results in a new light and open the way for researchers to use these instruments in the future. We hope that, by providing this detailed report of the process and demonstrating an effective use of content validation procedures, we will encourage other researchers in knowledge management and management information system (MIS) in general to undertake similar procedures in their projects, which will increase the number and quality of well-validated research instruments.

Conceptual models in MIS and in behavioural research in general typically involve constructs—variables that can only be measured indirectly via directly measurable indicators. For a given construct, multiple indicators capturing various aspects of the construct are typically employed (MacKenzie, Podsakoff, & Podsakoff, 2011; Straub, Boudreau, & Gefen, 2004). The indicator values are typically assessed in surveys via survey items. In order to test a conceptual model, a valid survey instrument is needed, one that measures what it is supposed to measure (DeVellis, 2003). Straub et al. (2004) argue that the instruments used to test a model need to be validated prior to testing the model to ensure that the findings and interpretations are based on valid data (i.e., data that have been collected using instruments with established validity).

Content validity refers to "the degree to which items in an instrument reflect the content universe to which the instrument will be generalized" (Straub et al., 2004, p. 424). For an instrument measuring a construct to be content valid, its indicators need to correctly capture all of the construct's important aspects and the questionnaire items must actually represent the content domain of their concepts (Haynes, Richard, & Kubany, 1995). Studies relying on an instrument lacking content validity may result in incorrect conclusions about relationships between constructs because important aspects of the constructs may be either lacking or be misrepresented. Thus, one could argue that, if a measure that does not obtain satisfactory evidence of construct validity, assessing a model need not proceed further (Scheirishim, Powers, Scandura, Gardiner, & Lankau, 1993). This also reflects Gable's (1986) view that "content validation should receive the highest priority during the process of instrument development" (p. 72). Straub et al. (2004) assert that validating one's instrument is a critical step before testing the conceptual model because the rigor of findings and interpretations is based on the instrument's foundations that are used to gather the data (Straub et al., 2004).

CONTRIBUTION

This study describes a quantitative approach to content validity to investigate KMS success factors for healthcare. It demonstrates the practicality of Lawshe's technique for content validity assessment when developing survey instruments. The descriptions of procedures for content validity assessment and the analysis of results using Lawshe's technique enhances IS researchers' understanding in creating good content for their measures. The importance of content validity is emphasized in the IS literature and yet few studies are reported. This study adds to the current literature on content validity assessment by providing an opportunity for more IS researchers to consider a quantitative approach for their content validity assessment in their scale-development process.

The IS literature suggests that it is an important and highly recommended practice to conduct content validity when developing a new instrument and also when applying existing scales to examine any new object (Straub et al., 2004).

Over the years, researchers have tried to establish methods for determining content validity. Previous studies have reported achieving content validity through a comprehensive review of the literature (e.g., Gold, Malhotra, & Segars, 2001; Kankanhalli, Tan, & Wei, 2005) and some (e.g., Bhatt, Gupta, & Kitchens, 2005; Janz & Prasarnphanich, 2003) through informal assessment by experts (e.g., Haynes et al., 1995; Tojib & Sugianto, 2006). Subject matter experts (SMEs) have an up-to-date knowledge of their content domain because they are professionally involved in growing, maintaining, and distributing such knowledge. Therefore, the involvement of SMEs in the content-validity process may provide useful insights into the completeness and appropriateness of a one's items, and their judgment is important for researchers to justify the content validity of their instruments (Lynn, 1986). However, in studies that involve a large pool of items, interpreting results can be difficult. Therefore, methods have been proposed to quantify the analysis (Cohen, 1968; Lawshe, 1975; Lynn, 1986) by conducting a quantitative assessment (Straub et al., 2004).

Lawshe (1975) has proposed one of the most commonly used approaches for qualitatively assessing construct validity using a group of SMEs. Wilson, Pan, and Schumsky (2012) overview the uses of the Lawshe's approach to assessing content validity in social sciences and, in this way, demonstrate its broad acceptance. Moreover, Lewis, Templeton, and Byrd's (2005) study of information resource management, which relied on the same analytical process as Lawshe initially suggests, provides an example of the use of this technique from the MIS field.

In our project, we investigated KMS success factors in healthcare organizations. To do so, we developed a self-administered questionnaire composed of multiple constructs and items. To develop the theoretical definition of the constructs we examined, we needed to understand the phenomenon to be investigated, which we developed by thoroughly review the literature. We adopted the items from existing instruments, which we found mainly in studies of business organizations (see Table 1). We reworded the instruments to fit the healthcare context. Although we used existing instruments, we did not know whether these instruments still accurately represented the constructs when used in a healthcare context. As such, we made content validity a primary concern when developing the instruments to ensure that the reworded items measure perceptions of KMS success in healthcare.

While there has been an increase in a number of KMS studies using survey instruments, there is little evidence that the content validity assessment of the instruments has been undertaken. For example, previous KMS studies (e.g., Wu & Wang, 2006; Kulkarni et al., 2007; Hwang, Chang, Chen, & Wu, 2008, Halawi et al., 2007) have developed instruments based on reviewing literature without reporting any procedures for content validity assessment.

Many of these previous studies of KMS success report assessing various types of instrument validity, such as convergent validity (which ensures that all reflective indicators of a given construct behave as if they reflect a construct in common) and discriminant validity (which ensures that all indicators of a given construct correlate with the construct the researcher intends to measure more strongly than with any other constructs in the model) (for an in-depth introduction of different types of validity, refer to MacKenzie et al. (2011) and Straub et al. (2004)). Nonetheless, no previous study of KMS success reports using a quantitative approach to content validity that relies on experts' judgments. Most previous studies report achieving content validity through a comprehensive review of literature (e.g., Gold et al., 2001; Kankanhalli et al., 2005) and some (e.g., Bhatt et al., 2005; Janz & Prasarnphanich, 2003) through informal assessment by experts.

Therefore, we validated our instrument using Lawshe's (1975) technique of evaluating the content validity ratio (CVR) to quantify the degree of consensus among a panel of experts. Although a few IS studies have previously used Lawshe's technique (e.g., Dwivedi, Choudrie, & Brinkman, 2006; Gable, Sedera, & Chan, 2008), these studies did not establish a series of steps to provide guidance in assessing the content validity of both individual items and measures.

In this paper, we demonstrate to KMS researchers a valuable and practical approach for assessing the content validation of instruments used in measuring KMS success factors in a healthcare context. In particular, we demonstrate how content validity procedures can be effectively used to determine content validity for multiple measures for single constructs through a series of steps using Lawshe's (1975) technique. This paper will be of interest to researchers and practitioners planning a content validity study, and to KMS researchers intending to reuse the measures we validate. Finally, we also outline the quantitative approach to assessing content validity, describes how we implemented the approach in this study, and present the results of the content validity assessment for the KMS success model's constructs.

BACKGROUND TO CONTENT VALIDITY

We can define content validity as the degree to which items of an assessment instrument are relevant and constitute an adequate operational definition of a construct (Adcock & Collier, 2001; Schriesheim et al., 1993; Waltz, Strickland, & Lenz, 2005; Wynd, Schmidt, & Schaefer, 2003). Content validity assesses which set of items are relevant and essential to represent the sample of the full domain of content. That is, an evaluation instrument with good content validity should include only items that are relevant and essential and that adequately cover the topics that have been defined as the relevant dimensions.

Information systems (IS) researchers have repeatedly raised the issue of instrument validation, particularly content validity, which was reportedly infrequently assessed in IS research (Boudreau, Gefen, & Straub, 2001; Straub, 1989; Straub et al., 2004; MacKenzie et al., 2011). In their 2001 assessment of instrument-validation practices in IS, Boudreau et al. (2001) indicate that only 23 percent of the paper they sampled examined content validity. In research that uses questionnaires as a method for collecting data, instruments are developed using a theoretical base, and items are generated based on existing instruments or new items are produced based on the researchers' understanding of the concepts suggested by theory. At the initial stage of an instrument's development, content validity is the primary concern that must be assessed immediately after items have been developed (Schriesheim et al., 1993). This may avoid problems associated with incomplete or biased measures which may result in researchers drawing wrong conclusions that may influence future decisions (McKenzie et al., 1999). Therefore, it is critical to investigate content validity prior to examining other types of validity to ensure the constructs are measured accurately (Hinkin, 1995; Lynn, 1986).

Haynes et al. (1995) emphasize that instruments lacking content validity may be valid in other ways. For example, accurate predictions may be possible using models with measures lacking content validity; lack of content validity does not prevent the measure from having predictive validity. Nonetheless, the interpretations of such accurate predictions are bound to be incorrect because interpretations are based on construct meanings, which are not correctly evaluated by instruments that lack content validity. Therefore, for research to result in a valid understanding of the social world, it is crucial that the instruments used are content valid.

Lynn (1986) suggests that content validity can be established via applying a two-stage process: development and judgement. The first stage (development) involves identifying the domain, generating the items, and formulating the instruments (Carmines & Zeller, 1979). Identifying the domain is an initial step to conceptually define the construct by reviewing the literature, which is followed by generating a set of items that are later arranged in a suitable sequence for the next stage of preparation. The second stage (judgment, which is the primary goal of content validation), involves asking several experts to evaluate the validity of individual items and the whole instrument. As a whole, this process helps a researcher retain the best items which are believed to adequately measure a desired content domain (Grant & Davis, 1997).

Some IS studies (e.g., Wang, 2003; Bhattacherjee, 2002) validated their contents by relying only on a comprehensive literature review; that is, they derived most of their items from an extensive literature review and existing instruments without having a panel of experts empirically assess them. Although the existing instruments have been validated, research recommends revalidating the instruments by consulting a panel of experts (Lynn, 1986).The experts will provide their opinion and constructive feedback about the quality of the items and their relevance to the targeted construct. Thus, this process will increase the researchers' confidence that the generated items adequately measure the construct. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999) suggest that the most common method of providing evidence of content validity for any test or assessment is to have content-area experts rate the degree to which each test item represents the objective or domain. Although most researchers have long acknowledged the importance of expert judgment to achieve content validity (Lawshe, 1975; Lynn, 1986; Polit & Beck, 2006), it is still a relatively unexplored area of validation in IS (Boudreau et al., 2001).

Experts' judgment can be assessed using qualitative or quantitative approaches (Haynes et al., 1995). Qualitative approaches involve a panel of experts who subjectively review the items by providing comments, ideas, or feedback regarding the items. This process does not involve statistical calculation. On the other hand, quantitative approaches seek a panel of experts' judgment about degree of each item's relevance to the construct. This approach relies on statistical analysis, which informs a researcher's final decisions about whether or not to retain the items.

Content validity assessments via literature review or via informal consultations with experts rely to a large extent on researchers' subjective judgments, which can result in biased outcomes. Moreover, because such assessments are unstructured, the processes involved may be difficult to reproduce. One can argue that informal approaches to assessing content validity have the advantage of offering researchers with maximal freedom to reshape the measures under assessment based on the insights they gain. Nonetheless, approaches that are both systematic (involve clearly specified, repeatable steps) and quantitative (rely on statistical tests rather than on subjective judgment) result in content validity tests that are objective and that can be independently verified. Therefore, in their content-validation guidelines developed based on critically reviewing existing practices, Haynes et al. (1995) emphasize the need to use systematic approaches and to quantify judgments, and MacKenzie et al. (2011), in their recently published guidelines for validation in MIS and behavioural research, suggest that researchers use a quantitative approach to testing content validity. Clearly, applying systematic, quantitative content validity tests to confirm the content validity of measures used in a research stream contributes to further validating the studies belonging to the research stream and is particularly important when no such tests have previously been conducted.

Several approaches have been proposed for quantitatively assessing content validity. Many of these approaches have focused on examining the consistency or consensus of a panel experts' evaluations (Polit, Beck, & Owen, 2007). Lindell and Brandt (1999) summarize the most common consensus methods to evaluate inter-rater agreement for content validity purposes proposed in psychology. One problem of such inter-rater methods is that they often do not allow for the possibility of experts to agree by chance, an issue which can be overcome be employing more-complex analyses such as Cohen's kappa (κ) (Cohen, 1968). The main problems with assessments based on inter-rater agreement are their computational complexity and the difficulty of interpreting their results because they focus on inter-rater agreement in general rather than on the specific issue of agreement that an item is "essential". Anderson and Gerbing's (1991) SA index approach involves experts assigning items to constructs, with the content validity judged based on the proportion of experts assigning the item to its intended construct.

Thus, we see that Anderson and Gerbing's (1991) SA approach, Schriesheim et al.'s (1993) approach, and Hinkin and Tracey's (1999) approach involve experts rating each item many times for each construct in a model with the resulting data factor-analysed. The approaches involving forcing experts to assign items to constructs have the drawback of the forced choice not reflecting the extent of certainty (Hinkin & Tracey, 1999). Moreover, the approaches requiring experts to rate each item for every construct in a model requires experts to expend excessive amounts of time and cognitive effort. It is likely that the value of additional data thus generated is offset by lower response rates, more missing data, and by experts paying less attention when rating the items that they do rate.

In our study, we followed Wilson et al.'s (2012) recommendation to use Lawshe (1975)'s approach, which involves experts rating individual items. The content validity ratio (CVR) is then calculated for each item, after which items are retained or removed based on their ratings. This approach is conceptually and computationally straightforward. We discuss it in more detail in the following section.

Content Validity Ratio

Lawshe (1975) developed one widely used method for measuring content validity, the CVR, which is used to measure items' content validity. Lawshe proposes a quantitative assessment of content validity, which employs a panel of subject matter experts (SMEs) to rate the degree to which each test item represents the objective or domain. In this approach, the SMEs are asked to evaluate the relevance of each of the construct on a three-point scale: "1 = not relevant", "2 = important (but not essential)", and "3 = essential".

The CVR is calculated based on the formula that Lawshe (1975) developed: $CVR = (2N_e / N) - 1$, where CVR = content validity ratio, N_e = number of SMEs indicating "essential", and N = total numbers of SMEs. The outcome of this formula is that:

- When all say "essential", the CVR is 1.00 (100% agreement)
- When the number saying "essential" is more than half (> 50%), but less than all (< 100%), the CVR is between zero and 0.99, and
- When fewer than half (< 50%) say "essential", the CVR is negative.

According to Lawshe (1975), if more than half the panellists indicate that an item is essential, that item has at least some content validity. Greater levels of content validity exist as larger numbers of panellists agree that a particular item is essential. An item would be considered potentially useful if its CVR fell in the range $0.0 < CVR <= 1.00$, which indicates that the number of panellists indicating "essential" is at least more than 50%. There is still the possibility that an item could be assigned a positive CVR value based purely on chance. Lawshe (1975) has established minimum CVRs for varying panel sizes based on a one tailed test at $p < 0.05$ significance level to establish the level at which a CVR is unlikely to be the result of chance.

Although Lawshe (1975) only utilizes the "essential" response category in computing the CVR, Lewis et al. (2005) employs a less-stringent criterion to compute CVR. They argue that responses in both the categories of "important (but not essential)" and "essential" should be utilized because both are positive indicators of the item's relevance to a construct.

In a previous study, Lyn (1986) created a similar table to Lawshe (1975); however, their approaches are not the same: Lyn uses normal distribution, which results in discrepancy. Her content validity index (CVI) procedure, which is determined by the proportion of experts who rate it as content valid, does not work for a small number of experts (when N <= 10). Lawshe's technique is better and more correct because he uses binomial distribution and his work can be used for a small number of experts. Lawshe's technique of using a table for determining a cutoff value provides an easy-to-compute method for quantification and significance testing that is more straightforward and user friendly.

Subject Matter Experts (SMEs)

Selecting the right experts, known as subject matter experts (SMEs), may influence the soundness of the validation process and can also be helpful in determining whether the measure is well constructed and suitable for psychometric testing (Davis, 1992). In his study, Davis suggests that expert panels should be comprised of individuals who have achieve professional certification in a related area, who have presented professional papers on the topic area, or who have initiated research in the topic area. Consistent with Davis (1992), Rubio, Berg-Weger, Tebb, Lee, and Rauch (2003) recommend that the experts (who they refer to as content eperts) should be professionals who have published or worked in the field. Rubio et al. (2003) also recommend using lay experts, who represent the population for whom the measure is being developed.

In terms of the required number of experts, Lynn (1986) recommends a minimum of three, while others (Gable & Wolf, 1993; Waltz et al., 2005) suggest a range of from two to twenty. Waltz et al. suggest at least two reviewers in the content area to be measured and one reviewer in instrument construction. Rubio et al. (2003) suggest using three experts for the two different groups that he identifies (i.e., professionals and lay experts). This number can range up to ten, which yields an expert sample size of six to twenty. With large sample sizes, more information can be generated about the measure. To increase the chance of identifying local colloquial terms inappropriate for an instrument, selecting members from different geographic locations is recommended (Grant, Kinney, & Guzzetta, 1990). However, the final decision for selecting the number of experts depends on experts' required expertise level and knowledge diversity (Grant & Davis, 1997).

METHODS

In our study, we derived the instrument development from an initial pool of items that we identified by comprehensively reviewing relevant KMS literature in business organizations and healthcare. We modified items adopted from studies in business organizations to suit the healthcare context. We used the DeLone and McLean (1992) IS success model as the basis for developing our model. We took or based the items for the constructs (perceived usefulness of KMS, user satisfaction, knowledge content quality, and KM system quality) from or the existing items in previous studies of DeLone and McLean's (1992) IS success model. We developed the items for the constructs KMS use for sharing and KMS use for retrieval based on reviewing the literature under Doll and Torkzadeh's (1998) guidelines. We took the items for the constructs incentive, subjective norm, perceived security, and culture of sharing from previous studies. The final model consisted of eleven constructs with 75 items. Table 1 summarizes the constructs appearing in our model, defines them, and lists their source.

Since this study reuses these items and constructs in a context that is quite different from their previous application, using existing instruments may not ensure that items are still valid. As such, we re-establish content validity in the healthcare context to ensure that the items represent the constructs and that each individual item measures what it is intended to measure.

After we generated the initial items, we designed the content validity study based on McKenzie, Wood, and Kotecki's (1999) guidelines and previous studies (Lyn, 1986; Davis, 1992; Rubio et al., 2003). Figure 1 outlines our procedure for conducting a content validity study using the quantitative method.

Table 1: Construct Definitions

Construct	Definition	Operationalization
KMS use for sharing	The extent to which KMS is being used for sharing.	Doll & Torkzadeh (1998)
KMS use for retrieval	The extent to which KMS is being used for retrieval.	Doll & Torkzadeh (1998)
Perceived usefulness of KMS	An employee’s perceptions of using KMS in the enhancement of his or her job performance.	Kulkarni et al. (2007)
User satisfaction	An employee’s attitude towards overall capabilities of the KMS in his or her organization.	Kulkarni et al. (2007), Wu & Wang (2006)
Knowledge content quality	How an employee perceives the quality of knowledge content in his or her organisation.	Halawi (2007), Kulkarni et al. (2007), Wu & Wang (2006)
KM system quality	The perceptions of an employee of the overall quality of the KMS as information systems.	Kulkarni et al. (2007), Wu & Wang (2006)
Leadership	The commitment of leaders with respect to knowledge management (KM) and their support for and encouragement of employees to share knowledge via KM.	Ash (1997), Guimaras, Igbaria, & Lu (1992), Kulkarni et al. (2007)
Incentive	The acknowledgement and recognition of knowledge sharing by employees.	Bock, Zmud, Kim, & Lee (2005), Lin (2007), Kulkarni et al. (2007)
Culture of sharing	The shared attitudes, values and practices of the employees in the organization with regard to knowledge sharing.	Kulkarni et al. (2007), Park, Ribier, & Schulte (2004)
Subjective norm	Perceptions of an employee regarding the peer pressure to share knowledge.	Ryu, Ho, & Han (2003)
Perceived security	The extent to which physicians believe that the KMS is secure for transmitting knowledge and trust that the knowledge shared is well-protected.	Fang, Chan, Brzezinski, & Xu (2006), Yenisey, Ozok, & Salvendy (2005)

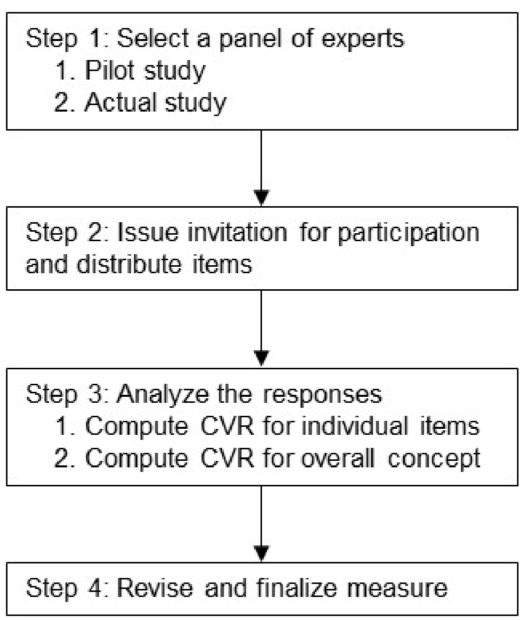

Figure 1: Procedures for conducting Expert Judgment

Step 1: Select a Panel of Experts

Following Davis's (1992) and Rubio et al.'s (2003) guidelines, we chose experts from both industry and academia who were well versed in the content area and who were knowledgeable about the development of survey measures. We developed our instruments to test the factors that influence the success of knowledge management systems in healthcare. Therefore, we chose experts who:

- hade working experience in the area of information systems and healthcare
- hade published papers in the concepts, theories, and implementation of KMS, and
- hade published papers in IT related to health.

The content validity process had two main phases: a pilot study and the actual study. For the pilot study, we used a convenience sample of five researchers; four were experts in IT-related health, and one was an expert in knowledge management. These experts had also published papers in peer-reviewed conferences. For actual study, we targeted twenty-five participants. We randomly chose twenty from the list of paper presenters in Health Informatics New Zealand Conference and five from the list of researchers who had published papers in the area of KMS in healthcare and KMS success models in peer-reviewed conferences and top-ranking journals such as *Information and Management* and *MIS Quarterly*. To help ensure our findings' generalizability and to help to identify local colloquialisms, we followed Grant et al.'s (1990) recommendations and chose participants from different countries: New Zealand, the USA, and Malaysia.

We approached a large number of experts to increase the chance of gaining an adequate number of responses because we assumed that not all experts would be willing to participate. We also considered the number we would need to meet the various minimum content validity ratio (CVR) values that satisfy the five percent probability level for statistical significance as Lawshe (1975) suggests. The greater the number of SMEs, the lower the minimum CVR required to meet the $p<0.05$ level.

Step 2: Issue Invitation for Participation and Distribute Items

For the pilot study, we sent five experts hard copies of the questionnaire. For the actual study, we approached twenty-five experts via an email invitation, which included a link to the URL of the response form (see Appendix 1). We used an online survey because such surveys are easily accessed and because our sample included experts from different geographic locations. We included a cover letter in the email to solicit panel members' participation. The cover letter explained the study's purpose, the reason we selected the expert, and the expert's role in the study should they agree to participate.

The survey formed provided instructions about how to use it, a description of the measure and its scoring, and an explanation of the form. We used both quantitative and qualitative reviews simultaneously as many researchers suggest,; this approach differs from the one McKenzie et al. (1999) suggests: that the qualitative and quantitative reviews be done separately.

For the quantitative review, we asked the respondents to evaluate an item's relevance and importance to the content domain, as described in the operational definition, by providing their rating for each item. We followed the criteria that Lawshe (1975) suggests to evaluate the measure using a scale from 1 to 3: 1) not relevant, 2) important (but not essential), and 3) essential. The last question asked the SMEs to rate the overall construct using the same scale points.

For the reviews' qualitative component, we asked the SMEs to provide feedback on overall items or to suggest any new items in the space provided.

Step 3: Analyse the responses

Compute CVR for individual items

From the total of 30 experts who we approached, we had 17 respondents. Five respondents were from the pilot study and 12 were from the actual study. One respondent did not rate the items, but provided comments on how to revise the measure (for a response rate of approximately 53%). As such, the total number of SMEs who responded and completed the form was 16 (N=16); however, some experts did not rate some of the items and, in some cases, they provided only 14 or 15 evaluations. From these data, we computed a content validity ratio (CVR) for each item. Having a panel of 16 experts was considered adequate because it is in the range that previous studies recommend. Furthermore, a range of between 10 and 20 SMEs allows greater consistency of response (Gable, 1986).

Although Lawshe (1975) only utilizes the "essential" response category in computing the CVR, we employed a less-stringent criterion to compute CVR as Lewis et al. (2005) suggest. We used responses in both the categories of "important (but not essential)" and "essential" because both were positive indicators of the items' relevance to KMS.

We did not use responses that did not provide a rating on a given item in the calculation of the CVR for that item. A positive value for the CVR indicates that more than 50 percent of the panelists rated the item as either “essential” or “important”.

As the number of respondents ranged from 14 to 16, we set our acceptance criteria at the more restrictive value that corresponds to a panel of 14 experts. Using the values Lawshe (1975) provides, we selected the minimum CVR value of 0.51 as our acceptance criteria. This means that, for a panel of 14 or more experts, there is only a 5 percent chance that enough experts would rate the item as content valid (i.e., a CVR value greater than 0.51) by chance. Of the 75 items, all items scored a positive CVR and 70 items (93%) had statistically significant CVR values (CVR > .51; p < .05).

Compute CVR for overall construct

Table 2 lists all the constructs and gives the total number of items evaluated and the number of items that gained statistically significant CVR values. We also asked the experts to rate the importance of each construct as a whole, and the table also presents the value of these CVRs. The CVR value for ten of the eleven constructs fell between the maximum value of 1.00 and minimum value of 0.63 and exceeded the 0.05 level of statistical significance, which suggests that these constructs are important for the KMS success model, although some of the items were rated with relatively low CVR. The incentive construct had a CVR of 0.27, which is below the threshold of chance but still indicates that the majority of experts considered it important.

The results of the content validity assessment illustrate that, overall, the constructs used for the KMS success model for healthcare possessed a high level of content validity and that most of the items were representative of the construct universe.

Table 2: The CVR Value for Each Construct

Constructs	Total items	Significant items	CVR of construct
KMS use for sharing	6	6	1.00
KMS use for retrieval	4	4	1.00
Perceived usefulness of KMS	7	7	1.00
User satisfaction	6	6	.81
Knowledge content quality	9	9	.63
KMS system quality	9	9	1.00
Leadership	7	7	1.00
Incentive*	5	1	.27
Culture of sharing	9	9	1.00
Subjective norm	4	4	1.00
Perceived security*	9	8	.63
Total number of items	75	70	

*Indicates constructs with some items that were not significant

Step 4: Revise and finalize measure

Only five items, of which 4 were from incentive construct and one was from perceived security construct, did not gain a statistically significant CVR value, although all gained a positive value. Table 3 presents the CVR value for each item in the constructs incentive and perceived security. We carefully reviewed these items and considered whether to revise or drop each one. In addition, we reviewed all items to consider the feedback from SMEs in relation to issues such as the clarity of the questionnaire’s structure and the wording of the items.

The non-significant results for the construct incentive and for the first four items from the construct incentive (see Table 3) suggests that there was only limited support from the SMEs for including these items, although positive CVR values indicate that the majority of the SMEs considered them important. These findings indicate that, even though these items may be suitable for measuring incentive in business organizations, they may be less appropriate in other contexts, such as healthcare. This may be because the nature of healthcare professionals’ job is to share knowledge to benefit their patients regardless of the incentives. In general in healthcare, where incentives exist, they are normally used in an attempt to increase physicians’ use of evidence-based treatments or to stimulate health professionals to change their clinical behavior with respect to preventive, diagnostic, and treatment decisions (Flodgren et al., 2011; Diamond & Sanjay, 2009) rather than to share knowledge.

We did not find the item from the construct perceived security that reads as "I believe that KMS has the mechanisms to protect knowledge from being stolen" to be significant. This suggests that this item is not a good measure for representing the security of knowledge. We had adapted the item from previous studies in which the construct perceived security (Fang et al., 2006; Salisbury, Pearson, Pearson, & Miller, 2001; Yenisey et al., 2005) was used in the context of security of online transactions, and our findings suggest that some of the SMEs did not feel that this item was as applicable in the new context.

Table 3: Original Items for the Constructs Incentive and Perceived Security

Constructs	Items	CVR
Incentive	I will receive financial incentives (e.g., higher bonus, higher salary) in return for my knowledge sharing.	0.07 *
	I will receive increased promotion opportunities in return for my knowledge sharing.	0.38 *
	I will receive increased job security in return for my knowledge sharing.	0.25 *
	Knowledge sharing is built into and monitored within the appraisal system.	0.43 *
	Generally, individuals are rewarded for teamwork.	0.62
Perceived security	I believe that knowledge I share will not be modified by the inappropriate parties.	0.87
	I believe that knowledge I share will only be accessed by authorized users.	0.88
	I believe that knowledge I share will be available to the right people.	0.88
	I believe that people in my organization do not use unauthorized knowledge.	0.73
	I believe that people in my organization use other's knowledge appropriately.	1.00
	I believe that KMS has the mechanisms to avoid the loss of critical knowledge.	0.88
	I believe that KMS has the mechanisms to protect knowledge from being stolen.	0.47 *
	In my opinion, the top management in my organization is entirely committed to security.	0.87
	Overall, I have confidence in knowledge sharing via KMS.	1.00

* $p < 0.05$; items dropped in the final version of KMS success model questionnaire.

Researchers recommend a second round of content-validity evaluation if the initial evaluation results in substantial changes to the instrument. Due to time constraints, we did not undertake a second round of evaluations in this case.

CONCLUSION

Content validity is a crucial factor in instrument development and yet is infrequently assessed in instrument validation, especially in the area of information systems. This paper demonstrates how to conduct a content validity study by describing the process used to assess the validity of instruments designed to test a model of KMS success in healthcare. This study illustrates the empirical assessment of the content validity, which can be a useful guideline for researchers in developing their own instruments. Using a panel of experts provides researchers with valuable information to revise a measure. The content validity assessments reported in this study confirmed the content of measures for a KMS success model for healthcare. Further, by demonstrating the effective use of content validation for a KMS success model for healthcare, this study may promote the use of content validation in the broader contexts of KMS and health information systems research. The methodology we describe in this paper is relatively easy to implement and provides worthwhile feedback; we recommend that future empirical studies apply similar procedures to increase the confidence in their results.

REFERENCES

Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review, 95*(3), 529-546.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, *76*(5), 732-740.

Ash, J. (1997). Organizational factors that influence information technology diffusion in academic health sciences centers. *Journal of the American Medical Information Association, 4*(2), 102-111.

Bhatt, G. D., Gupta, J. N. D., & Kitchens, F. (2005). An exploratory study of groupware use in the knowledge management process. *The Journal of Enterprise Information Management, 18(*1), 28-46.

Bhattacherjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of Management Information Systems*, *19(*1), 211-242.

Bock, G.W., Zmud, R. W., Kim, Y. G., & Lee, J. N. (2005). Behavioral intention formation in knowledge sharing: Examining the role of extrinsic motivators, social-psychological forces, and organizational climate, *MIS Quarterly, 29*(1), 87-111.

Boudreau, M. C., Gefen, D., & Straub, D. W. (2001). Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly, 25*(1), 1-16.

Carmines, E. G., & Zeller, R. A. (Eds.). (1979). *Reliability and validity assessment*, Thousand Oaks, CA: Sage.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213.

Davis, L. (1992). Instrument review: Getting the most from your panel of experts. *Applied Nursing Research, 5*(4), 104–107.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: the quest for the dependent variable. *Information Systems Research*, *3*(1), 60-95.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Diamond, G. A., & Sanjay, K. (2009). Evidence-based financial incentives for healthcare reform putting it together. *Circulation: Cardiovascular quality and outcomes*, *2*(2), 134-140.

Doll, W. J. & Torkzadeh, G. (1998). Developing a multidimensional measure of system-use in an organizational context. *Information and Management*, *33*(4), 171-185.

Dwivedi, Y. K., Choudrie, J. & Brinkman, W. P. (2006). Development of a survey instrument to examine consumer adoption of broadband. *Industrial Management and Data Systems*, *106*(5), 700-718.

Fang, X., Chan, S., Brzezinski, J., & Xu, S. (2006). Moderating effects of task type on wireless technology acceptance. *Journal of Management Information Systems*. *22*(3), 123-157.

Flodgren, G, Eccles, M. P., Shepperd, S., Scott, A., Parmelli, E., & Beyer, F. R. (2011). An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Sys Ret*, *7*.

Gable, R. K. (1986). *Instrument development in the affective domain*, Boston: Kluwer-Nijhof.

Gable, R .K., & Wolf, J. W. (1993). *Instrument development in the affective domain: Measuring attitudes and values in corporate and school settings,* Boston: Kluwer Academic.

Gable, G. G., Sedera, D., & Chan, T. (2008). Re-conceptualizing information system success: The IS-impact measurement model. *Journal of the Association for Information Systems*, *9*(7), 1-32.

Gold, A. H., Malhotra, A., & Segars, A.H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185-214.

Guimaraes, T., Igbaria, M., & Lu, M-T. (1992). The determinants of DSS success: An integrated model. *Decision Sciences, 23*(2), 409-430.

Grant, J. S., & Davis, L. T. (1997). Selection and use of content experts in instrument development. *Research in Nursing and Health, 20*(3), 269–274.

Grant, J., Kinney, M., & Guzzetta, C. (1990). A methodology for validating nursing diagnoses. *Advances in Nursing Science, 12*(3), 65–74.

Halawi, L. A., McCarthy, R. V., & Aronson, J. E. (2007). An empirical investigation of knowledge management systems' success. *Journal of Computer Information Systems, 48*(2), 121-135.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238-247.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967-988.

Hinkin, T. R., & Tracey, J. B. (1999). An analysis of variance approach to content validation. *Organizational Research Methods*, *2*(2), 175-186.

Hwang, H.G., Chang, I., Chen, F. J., & Wu, S. Y. (2008). Investigation of the application of KMS for diseases classifications: A study in a Taiwanese hospital. *Expert Systems with Applications*, *34*(1), 725-733.

Janz, B. D., & Prasarnphanich, P. (2003). Understanding the antecedents of effective knowledge management: The importance of a knowledge-centered culture. *Decision Sciences, 34*(2), 351-384.

Jennex, M., & Olfman, L. (2005). Assessing knowledge management success. *International Journal of Knowledge Management, 1*(2), 33-49.

Kankanhalli, A., Tan, B. C. Y., & Wei, K. K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly, 29*(1), 113-143.

Kulkarni, U. R., Ravindran, S., & Freeze, R. (2007). A knowledge management success model: Theoretical development and empirical validation. *Journal of Management Information Systems,* 23(3), 309-347.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563-575.

Lewis, B. R., Templeton, G. F., & Byrd, T. A. (2005). A methodology for construct development in MIS research. *European Journal of Information Systems, 14*(4), 388-400.

Lin, H-F. (2007). Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of Information Science, 33*(2), 135-149.

Lindell, M. K., & Brandt, C. J. (1999). Assessing inter-rater agreement on the job relevance of a test: A comparison of the CVI, T, $r_{WG(J)}$, and $r^*_{WG(J)}$ indexes. *Journal of Applied Psychology*, *84*(4), 640–647.

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385.

McKenzie, J. F., Wood, M. L., & Kotecki, J. E. (1999). Establishing content validity: Using qualitative And quantitative steps. *American Journal of Health Behavior, 23*(4), 311-318.

MacKenzie, S. B., Podsakoff P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293-334.

Park, H., Ribier, V., & Schulte, W.D. (2004). Critical attributes of organizational culture that promote knowledge management technology implementation success. *Journal of Knowledge Management, 8*(3), 106-117.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing and Health, 29*(5), 489–497.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in nursing and health, 30*(4), 459-467.

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work. *Social Work Research, 27*(2), 94-104.

Ryu, S., Ho, S.H., & Han, I. (2003). Knowledge sharing behavior of physicians in hospitals. *Expert Systems with Applications, 25*(1), 113-122.

Salisbury, D. W., Pearson, R. A., Pearson, A. W., & Miller, D. W. (2001). Perceived security and World Wide Web purchase intention. *Industrial Management and Data Systems, 101*(4), 165-176.

Schriesheim, C. A., Powers, K. J., Scandura, T.A., Gardiner, C. C., & Lankau, J. (1993). Improving construct measurement in management research: Comments and a quantitative approach for assessing the theoretical content adequacy of paper-and-pencil survey-type instruments. *Journal of Management, 19*(2), 385-417.

Straub, D. W., (1989). Validating Instruments in MIS Research. *MIS Quarterly*, *13*(2), 147-170.

Straub, D. W., Boudreau, M. C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the AIS, 13*(1), 380-427.

Tojib, D. R., & Sugianto, L.-F., (2006). Content validity of instruments in IS Research. *Journal of Information Technology Theory and Application, 8*(3), 31-56.

Waltz, C. F., Strickland, O. L., & Lenz E. R. (2005). *Measurement in Nursing and Health Research* (3rd ed.). New York: Springer.

Wang, Y. S. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information and Management*, *41*(1), 75-86.

Wilson, F. R., Pan W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counselling and Development*, *45*(3), 197-210.

Winkelman, W. J., & Choo, C. W. (2003). Provider-sponsored virtual communities for chronic patients - improving health outcomes through organizational patient-centred knowledge management. *Health Expectations*, *6*(4), 352-358.

Wu, J. H., & Wang, Y. M. (2006). Measuring KMS success: A respecification of the DeLone and McLean's model. *Information and Management*, *43*(6), 728-739.

Wynd, C. A., Schmidt, B., & Schaefer, M. A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research, 25*(5), 508-518.

Yenisey, M. M., Ozok, A. A., & Salvendy, G., (2005). Perceived security determinants in e-commerce among Turkish university students. *Behaviour and Information Technology, 24*(4), 259-274.

APPENDIX 1: COVER LETTER AND EXTRACT OF QUESTIONNAIRE

Dear ,

I noted that your presentation at HINZ2009 is related to the topic of my research, namely, the success of knowledge management systems in healthcare. As part of my PhD project I conducted a comprehensive literature review, exploring the use of information systems for knowledge management in healthcare (covering over 100 journal articles).

I wonder if you might be interested in acting as one of the experts in my content validity study. Your role would be to rate the importance of various determinants of knowledge management systems success in a model that I formulated based on the literature review.

As the model summarizes a large body of literature related to your interests, I am certain that you will find it to be quite informative.

You can rate the concepts online at
http://is-research.massey.ac.nz/kmsh (It should take about 10 minutes).

Please use "102" as the token number to participate in the survey.

As a participant, you will receive a report outlining the results of the content validity study (while preserving strict confidentiality) and summarizing the literature on which the content validity study instrument is based.

Your input will be highly crucial for the success of my research, as it will enable me to determine the right constructs for the research model and to justify their inclusion.

Attached is the information sheet which gives further details of my research.

If you know of anyone else who may be interested in being a part of this study, I would be grateful if you put me in contact with them.

Thank you for considering this. I look forward to hearing from you.

Yours Sincerely,

Nor'ashikin Ali Doctoral research student

ANSWER SHEET

We ask you, as an expert, to evaluate a questionnaire devoted to the determinants of Knowledge Management Systems success in healthcare. In the following, the text of the questionnaire that we ask you to evaluate is given in blue, while our instructions to you as an evaluator are given in black. No matter what we ask you to do explicitly, please feel free to also provide free-form comments by writing anywhere on this form.

A. KMS Use for Sharing

The extent to which KMS is being used to communicate with colleagues, participate in online discussion groups, discuss ideas/views/experiences, collaborate with peers, distribute knowledge, identify and locate people for knowledge and expertise.

		Please, evaluate the importance of each item for measuring the concept of "KMS use for sharing" by ticking (√) as appropriate		
	Item	**Not relevant**	**Important (But Not Essential)**	**Essential**
A1.	I use KMS to communicate knowledge with colleagues	☐	☐	☐
A2.	I use KMS to contribute ideas and/or feedbacks to the forum	☐	☐	☐
A3.	I use KMS to participate in online discussion groups	☐	☐	☐
A4.	I use KMS to discuss and/or exchange ideas/views/experiences with colleagues	☐	☐	☐
A5.	I use KMS to collaborate with colleagues	☐	☐	☐
A6.	I used KMS to distribute knowledge (e.g. news, memos, reports, presentation, organization policies)	☐	☐	☐

Please, suggest any other items that you think are important for measuring the construct of "KMS use for sharing", or further comment on the items we suggested.

____
____
____
____
____
____
____
____

ABOUT THE AUTHORS

Nor’ashikin Ali graduated from Massey University, New Zealand with a PhD in Information Systems. Her PhD work focused on knowledge management systems uccess model in healthcare. She is currently a senior lecturer at Universiti Tenaga Nasional, Malaysia, where she teaches knowledge management as her core subject. Her current research focuses on knowledge management in healthcare and Malaysian public sectors.

Alexei Tretiakov is a Senior Lecturer at Massey University’s School of Management. He has a PhD in Information Sciences from Tohoku University, Sendai, Japan. His current research interests are in information systems success.

Dick Whiddett graduated from Lancaster University, UK with a PhD in Computer Science and is currently a senior lecturer at Massey University, New Zealand where he teaches in the areas of Information Systems, Health Informatics and Health Service Management. Much of his research has examined the factors which influence the adoption of information technologies with a particular interest in electronic health records

JITTA

JOURNAL OF INFORMATION TECHNOLOGY THEORY AND APPLICATION

JITTA is a Publication of the Association for Information Systems
ISSN: 1532-3416